



PROJECT REPORT
TWO STEP COMPRESSION
USING HUFFMAN & DYNAMIC REPLACE WORDS ALGORITHM
ON HTML & TEXT FILE

Edward Christoforus Mo'a

03.02.0054

2008

FACULTY OF COMPUTER SCIENCE
SOEGIJAPRANATA CATHOLIC UNIVERSITY

JL. PAWIYATAN LUHUR IV/1, BENDAN DUWUR, SEMARANG 50234

Phone. 024-8441555 (hunting) Web: <http://www.unika.ac.id>

Email: ikom@unika.ac.id

APPROVAL AND RATIFICATION PAGE

PROJECT REPORT TWO STEP COMPRESSION USING HUFFMAN & DYNAMIC REPLACE WORDS ALGORITHM ON HTML & TEXT FILE

This project report already approved and ratified by Dean of Faculty of Computer Science and supervisor on

With the approval,

Examiner,

Examiner,

Suyanto EA, Ir.MSc
NIP. 058.1.1992.116

H.Marlon Leong, S.Kom, M.Kom,
NIP. 058.1.2007.273

Supervisor,

Dean of Faculty of Computer Science,

Ridwan Sanjaya., SE, S.Kom, MS.IEC
NIP. 058.1.2002.252

Ridwan Sanjaya., SE, S.Kom, MS.IEC
NIP. 058.1.2002.252

STATEMENT OF ORIGINALITY

I, the undersigned :

Name : Edward Christoforus Mo'a

ID : 03.02.0054

Hereby certify that this project was made by my self and not copy or plagiarizes from other people, except that in writing expressed to the other article.

If next time prove that this project was plagiarizes or copy the other, hence I readies receives a sanction.

Semarang, 11 August 2008

Edward Christoforus Mo'a
03.02.0054

FOREWORD

First of all, I am grateful to the Lord, Jesus Christ, who has been always with me in all my life, specially when I was doing this final project entitled “TWO STEP COMPRESSION USING HUFFMAN AND DYNAMIC REPLACE WORDS ALGORITHM”. I do feel His invisible helping hands in facing so many different difficulties and handicaps that has led me to the best solution.

There have been a lot of experiences I have got during my doing this final project, both good ones and bad ones. They all, however, have played an important role in making me to be better and to be closer to God, the Almighty.

As a matter of fact, I have been studying for five years in Faculty of Computer Science, and I feel I owe special thanks to lectures from whom I have got valuable guidance, directives and support, as well as the officials who have played a significant role in the success of my study. I also found out that I was not just learning about computer science but more than that I also learned a lot about how the life is that brings me to become a better person.

My special gratitude also goes to Mr. Ridwan Sanjaya, SE, Skom, MS IEC, my supervisor, who is good lecturer and has been so patient giving me guidance and helping me in the process of doing this final project.

I proudly dedicate this Final Project to my parents, my brother and sister, and all the members of my big family who have been always supporting me with deep love, continuous prayer and nice understanding.

I owe special thanks to friends who helped me in various ways to evade improper difficulties during my doing this final project who I cannot mention here one by one.

Last but not least, I would like to apologize for any mistake I have done in completing this project, and accordingly constructive critics and suggestions would be welcome.

Semarang, 11 August 2008

Edward Christoforus Mo'a

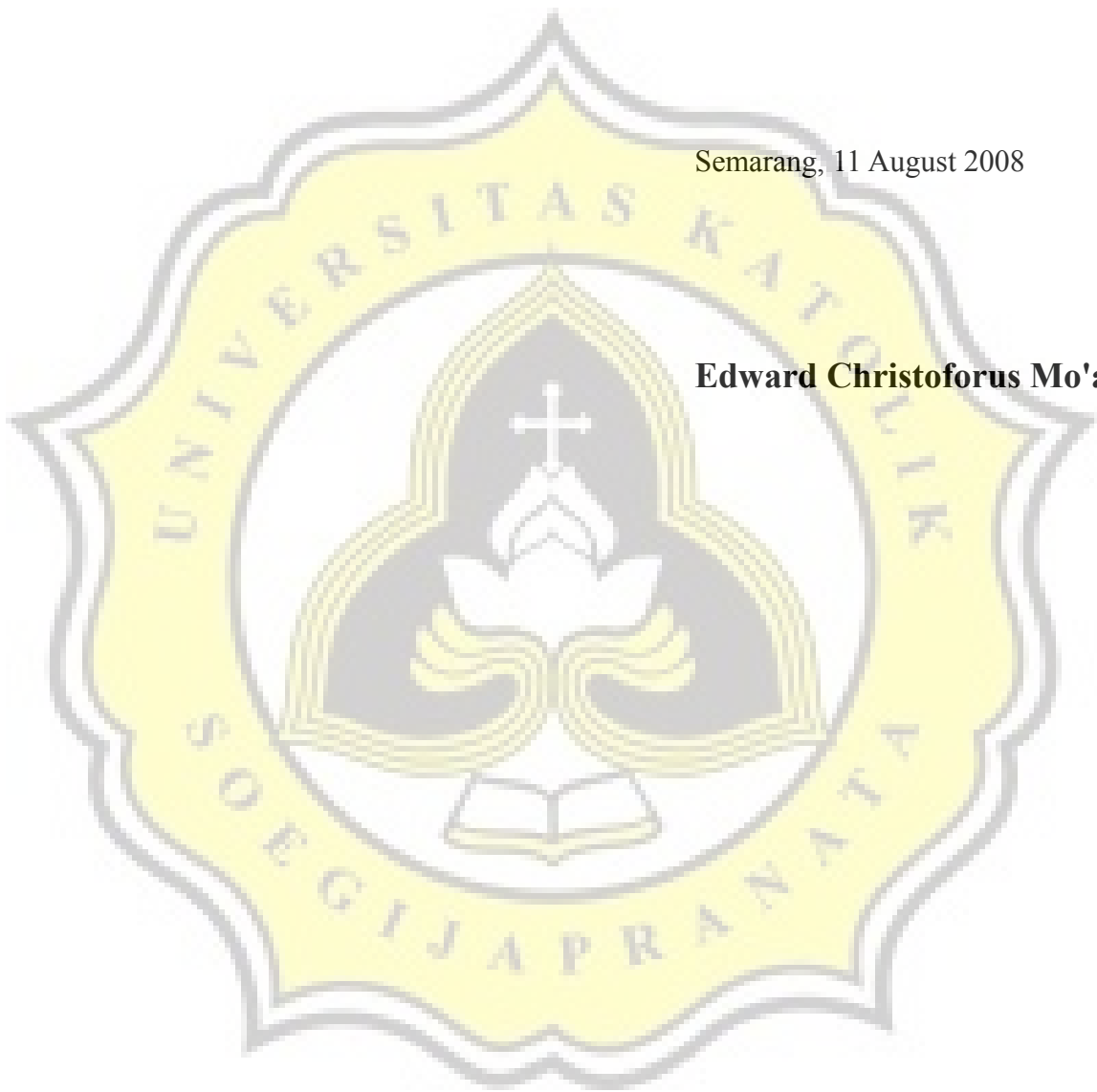


TABLE OF CONTENTS

APPROVAL AND RATIFICATION PAGE.....	ii
STATEMENT OF ORIGINALITY	iii
FOREWORD.....	iv
TABLE OF CONTENT.....	vi
TABLE OF FIGURE.....	ix
TABLE OF TABLES.....	x
ABSTRACT.....	xi
CHAPTER I INTRODUCTION.....	1
1.1. Background.....	1
1.2. Scope.....	3
1.3. The Formulation of Problem.....	3
1.4. The Purposes of this Final Project.....	3
1.5. Overview Documents.....	4
CHAPTER II LITERATURE STUDY	5
2.1 Huffman Algorithm	5
2.1.1 The process of building Huffman Tree.....	6
2.1.2 The Compression process using Huffman Algorithm... ..	12
2.1.3 The Decompression process using Huffman Algorithm.....	16
2.2 Dynamic Replace Words Algorithm.....	17
2.2.1 The Compression process using Dynamic Replace Words Algorithm.....	18
2.2.2 The Compression process using Dynamic Replace Words Algorithm.....	24
2.3 Tree Data Structure.....	26

CHAPTER III PLANNING.....	29
3.1. Research Methodology.....	29
3.2. Project Management.....	30
CHAPTER IV ANALYSIS AND DESIGN.....	31
4.1 Analysis.....	31
4.1.1 Functional Requirements.....	31
4.1.2 Non Functional Requirement.....	33
4.1.3 Constraints.....	33
4.2. Design.....	33
4.2.1 Use Case Diagram.....	35
4.2.2 Class Diagram.....	39
4.2.2.1 High Level Class Diagram.....	39
4.2.2.2 Low Level Class Diagram.....	40
CHAPTER V SOFTWARE IMPLEMENTATION & TESTING.....	44
5.1. Software Implementation.....	44
5.1.1 Output appearance of Two Step Compression Project.....	44
5.1.2 The Implementation of Huffman Algorithm in Two Step Compression Project.....	48
5.1.2.1 Implementation of compression process using Huffman Algorithm.....	48
5.1.2.2 Implementation of Decompression process using Huffman Algorithm.....	51
5.1.3 The Implementation of Dynamic Replace Words Algorithm in Two Step Compression Project.....	52

5.1.3.1	The Implementation of Compression process using Dynamic Replace Words Algorithm.....	53
5.1.3.2	The Implementation of Decompression process using Dynamic Replace Words Algorithm.....	55
5.2.	Testing.....	57

CHAPTER VI CONCLUSION & FURTHER RESEARCH....		58
6.1	Conclusion.....	58
6.2	Further research.....	61
References		62
APPENDIX Two Step Compression program code.....		63

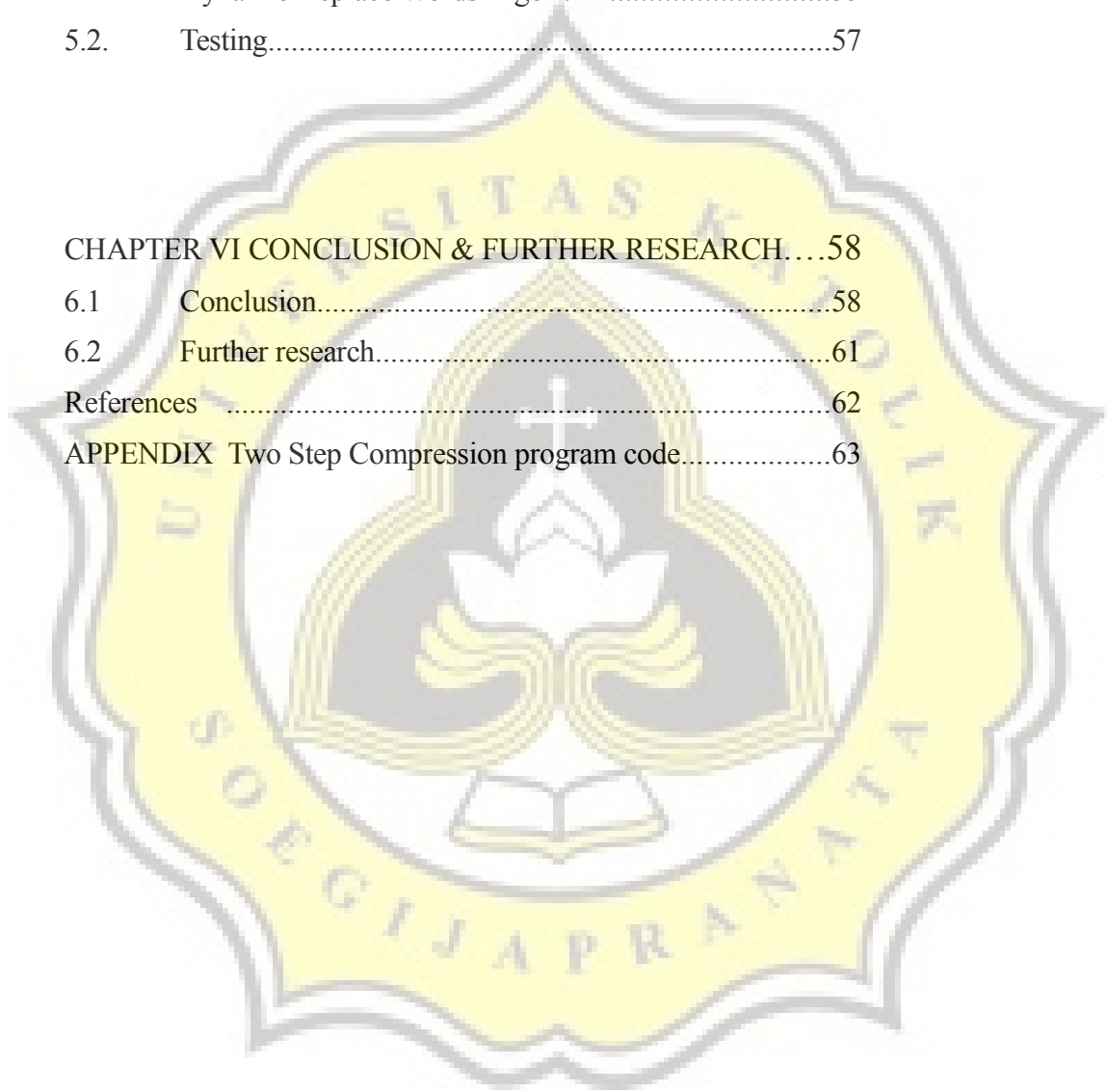
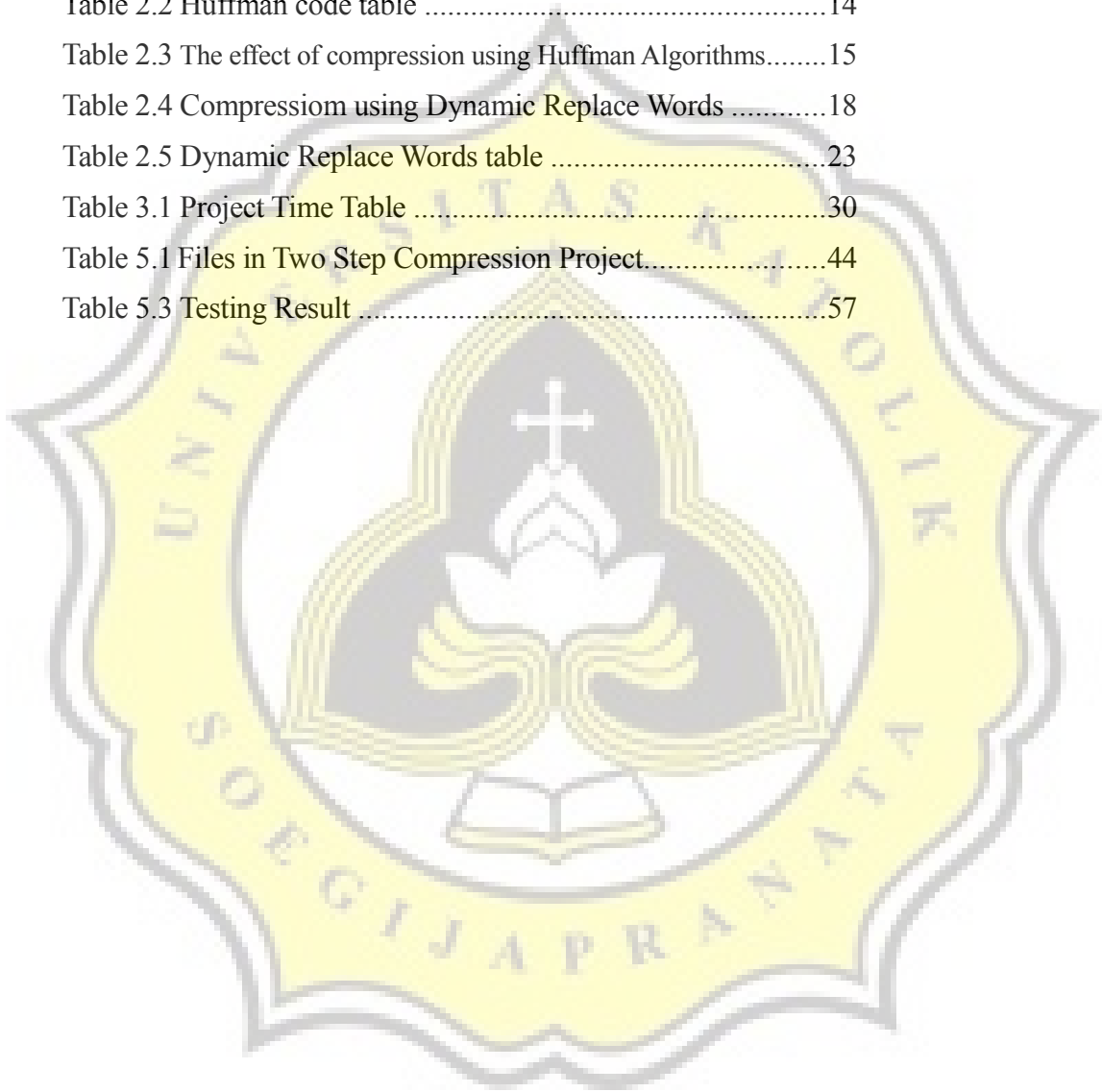


TABLE OF FIGURE

Figure 2.1 Growing the Huffman tree, Part 1.....	7
Figure 2.2 Growing the Huffman tree, Part 2	8
Figure 2.3 Growing the Huffman tree, Part 3	9
Figure 2.4 Growing the Huffman tree, Part 4	10
Figure 2.5 The final Huffman tree	11
Figure 2.6 Huffman encoding for character b.....	13
Figure 2.7 Seperate by space process.....	19
Figure 2.8 First node in Dynamic Replace Words Tree.....	20
Figure 2.9 Comparison process.....	21
Figure 2.10 Build the Dynamic Replace Words Tree, Part1.....	21
Figure 2.11 The Final Dynamic Replace Words Tree.....	22
Figure 2.12 Structure of Dynamic Replace Words file compress.....	24
Figure 2.13 Read Table process	25
Figure 2.14 Dynamic Replace Words decompress write original file process	26
Figure 4.1 The compression process chart.....	34
Figure 4.2 The Decompression process chart.....	35
Figure 4.3 Use case diagram Two Step Compression Project.....	36
Figure 4.4 High Level Class Diagram.....	39
Figure 5.1 Two Step Compression GUI appearance.....	45
Figure 5.2 Choosing file to be compressed.....	46
Figure 5.3 The compression process is successful.....	46
Figure 5.4 Choosing file to be decompressed.....	47
Figure 5.5 The decompression process is successful.....	47

TABLE OF TABLES

Table 2.1 Frequency Table.....	14
Table 2.2 Huffman code table	14
Table 2.3 The effect of compression using Huffman Algorithms.....	15
Table 2.4 Compression using Dynamic Replace Words	18
Table 2.5 Dynamic Replace Words table	23
Table 3.1 Project Time Table	30
Table 5.1 Files in Two Step Compression Project.....	44
Table 5.3 Testing Result	57



TWO STEP *COMPRESSION*

USING HUFFMAN ALGORITHM & DYNAMIC REPLACE WORDS

ON HTML & TEXT FILE

Edward Christoforus Mo'a (03.02.0054)

Faculty of Computer Science Soegijapranata Catholic University

edwardchristof@student.unika.ac.id

Abstract

*Nowadays a computer can handle a lot of different work, starting from data processing, digital features processing, multimedia, games, etc. The ability of the computer should be certainly balanced by a proper storage medium. Anyway, whatever the size of the storage of a medium, it will at last be full. Then what is the solution? *COMPRESSION* is one of the solutions! Because by making compression on the file which is going to be saved in the storage media, it is hoped to save disk space on the storage media.*

In this Final Project, the application is done in HTML and text file using Huffman Algorithm and Dynamic Replace Words Algorithm. The Data Structure tree is used in this application. Huffman Algorithm is suitable for this project because there is a frequency count of appearing data and the coding of data becomes smaller than it was before. While Tree is used to form dictionary file. The Replace character in this Project is found through a dynamic searching method and frequency count of appearing character. It is hoped through this method its compression scope is larger and the result of the compression is more efficient as well.

Keywords : Compression, Huffman Algorithm, Dynamic Replace Words Algorithm, HTML file, text file